

Fitting the Smile Revisited: A Least Squares Kernel Estimator for the Implied Volatility Surface

Matthias R. Fengler*

Institut für Ökonometrie und Statistik, Wirtschaftswissenschaftliche Fakultät
CASE – Center for Applied Statistics and Economics
Humboldt-Universität zu Berlin,
Spandauer Straße 1, 10178 Berlin, Germany
fengler@wiwi.hu-berlin.de
TEL ++49 30 2093 5654
FAX ++49 30 2093 5649

Qihua Wang[†]

Academy of Mathematics and System Science
Chinese Academy of Science
Beijing, 100080
P.R. China
qhwang@mail.amss.ac.cn

August 6, 2003

*I gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft and the Sonderforschungsbereich 373 “Quantifikation und Simulation ökonomischer Prozesse”.

[†]The second author was supported by the national natural scientific foundation (key grant: 10231030).

Abstract

Nonparametric methods for estimating the implied volatility surface or the implied volatility smile are very popular, since they do not impose a specific functional form on the estimate. Traditionally, these methods are two-step estimators. The first step requires to extract implied volatility data from observed option prices, in the second step the actual fitting algorithm is applied. These two-step estimators may be seriously biased when option prices are observed with measurement errors. Moreover, after the nonlinear transformation of the option prices the error distribution will be complicated and less tractable. In this study, we propose a one-step estimator for the implied volatility surface based on a least squares kernel smoother of the Black-Scholes formula. Consistency and the asymptotic distribution of the estimate are provided. We demonstrate the estimator using German DAX index option data to recover the smile and the implied volatility surface.

Keywords: implied volatility surface, smile, Black-Scholes formula, least squares kernel smoothing

1 Introduction

Functional flexibility is the key challenge for model building and model selection in quantitative finance: often it is difficult, or sometimes impossible to justify on theoretical grounds a specific parametric form of the economic relationship under investigation. Furthermore, in a dynamic context, the economic structure may be liable to sizable changes and considerable fluctuation. Thus, estimation techniques that do not impose any a priori restrictions on the estimate, such as non- and semiparametric methods, are increasingly popular in financial practice.

Recently in finance, the implied volatility surface (IVS) has attracted much attention. Since it is derived from option prices observed at a certain point in time via the Black-Scholes (BS) formula, it is a widely accepted state variable that reflects current market expectations and market sentiments in a forward looking and instantaneous manner. Traders continuously monitor and update the IVS they ‘trade on’. Pricing and risk management tools are fed with the IVS currently prevailing in the market. In the case of the IVS, however, model flexibility is a necessity rather than an option: from the BS theory, the IVS should be a flat and constant function across strike prices K and the term structure τ of the option’s time to maturity. Yet, as a matter of fact, one observes rich functional patterns fluctuating through time, see Figure (1) for illustration. This dependence, captured by the function $\hat{\sigma}_t : (K, \tau) \rightarrow \hat{\sigma}_t(K, \tau)$, is called IVS.

Parametric attempts to model the IVS along the strike profile, i.e. the ‘smile’, usually employ quadratic specifications, Shimko (1993), Ané and Geman (1999), and Tompkins (1999) among others. However, it seems that these approaches are not capable of capturing the salient features of IVS patterns, and hence estimates may be biased. In a comprehensive study, Dumas et al. (1998) document this for a variety of parametric choices of local volatility models. In an attempt to allow for more flexibility, Hafner and Wallmeier (2001) fit quadratic splines to the smile function. Aït-Sahalia and Lo (1998); Rosenberg (2000); Aït-Sahalia et al. (2001b); Cont and da Fonseca (2002); Fengler et al. (2003a) employ a Nadaraya-Watson estimator of the IVS function, and higher order local polynomial smoothing of the IVS is used in Rookley (1997). Aït-Sahalia et al. (2001a) discuss model selection between fully parametric, semi- and nonparametric IVS specifications and argue in favor of the latter approaches. In

IVS on Jan. 02, 2001

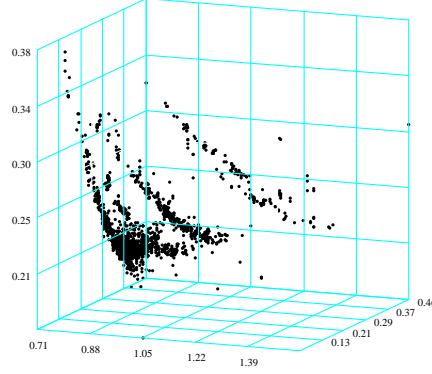


Figure 1: *IVS as observed on Jan. 02, 2001, calculated separately for each observed option price from the DAX Index options and discounted DAX Future prices. Lower front axis is moneyness $\kappa_t = K/(F_t e^{-r\tau})$ and lower right axis time to maturity measured in years corresponding to 17, 45, 73, and 164 days to expiry.*

combining techniques from functional data analysis and backfitting of generalized additive models, Fengler et al. (2003b) propose a modeling approach of the IVS that fits only in the local neighborhood of the observed design points of the IVS. An estimate of the IVS of the particular day is then given by a sum of smooth basis functions, whose weights may change over time. The decisive advantage of this estimator is in the fact that it is tailored to the degenerated, i.e. discrete, ‘string’ data structure, which can be seen in Figure (1).

All these approaches share in common that they are *two-step* estimators by nature: in a *first* step, implied volatilities are derived by equating the BS formula with observed market prices and solving for the diffusion coefficient, Manaster and Koehler (1982); in the *second* step the actual fitting algorithm is applied. These two-step estimators may be seriously biased, when option prices or other input parameters can only be observed with errors. Moreover, the nonlinear transformation of option prices makes the error distribution less tractable. Indeed, it has been conjectured that the presence of measurement errors can be of substantial impact, see Roll (1984), Harvey and Whaley (1991), and particularly Hentschel (2002) for

an extensive study on errors in implied volatility estimation and their possible magnitude. Potential error sources are the bid-ask bounce, nonsynchronous pricing, infrequent trading of index stocks, and finite quote precision.

Here, we propose a one-step procedure based on a least squares kernel estimator that directly takes option prices, say calls \tilde{C}_t , and other variables observed at time t as input parameters:

$$\hat{\sigma}(\kappa_t, \tau) = \arg \min_{\sigma} \sum_{i=1}^n \left\{ \tilde{C}_{t_i} - S_{t_i} C^{BS}(\cdot, \sigma) \right\}^2 W(\kappa_{t_i}) K_1 \left(\frac{\kappa_t - \kappa_{t_i}}{h_{1,n}} \right) K_2 \left(\frac{\tau - \tau_i}{h_{2,n}} \right),$$

where $C^{BS}(\cdot, \sigma)$ denotes the BS price for calls, $\kappa_t = K/S_t$ is moneyness, i.e. the rescaled strike dimension ‘strike by asset price’. $K_1(\cdot)$ and $K_2(\cdot)$ are kernel functions. $W(\cdot)$ denotes a uniformly continuous and bounded weight function, which allows for differential weights of observed option prices. These weights may be useful in the following respect: it is usually argued that in-the-money options contain a liquidity premium and should be incorporated to a lesser extent into the implied volatility estimate, or even excluded, Aït-Sahalia and Lo (1998); Skiadopoulos et al. (1999). This goal may be achieved in using the weight function $W(\kappa)$.

Our estimator can be interpreted as a *localized* version of estimation approaches in the early implied volatility literature, Latané and Rendelman (1976); Chiras and Manaster (1978); Schmalensee and Trippi (1978); Beckers (1981); Whaley (1982). These studies are little concerned in obtaining a fit of the IVS as such, but rather in recovering one single, good predictor of future stock price variability: e.g. Whaley (1982) minimizes an equally weighted sum of squared pricing errors. This line of literature, however, fails to explore the asymptotic properties of their estimators.

We provide asymptotic results, and computation of confidence intervals around the estimates is straight forward. Inspecting confidence intervals of the fitted smile is notoriously neglected in the implied volatility literature, even when traditional two-step estimators are used. In the computation of our confidence bands the nonlinear inversion of option prices to derive implied volatility is explicitly taken into account. Thus they are more accurate than those obtained from two-step estimators. Confidence bands are critically important to traders, since they help them to discern for a given IVS estimate, whether new observations constitute a significant change of the IVS. Furthermore, trading models may be constructed in a way

such that the generation of trading signals is based on the confidence of the current IVS estimate.

The paper is structured as follows: in Section (2) the estimator is given in detail, consistency and the asymptotic distribution theory is established. Section (3) presents IVS estimates on German DAX index data, Section (4) concludes. Proofs are given in an appendix.

2 Least Squares Kernel Smoothing of the IVS

European style calls and puts are contingent claims on an asset S_t (for simplicity, paying no dividends, here), which yield as pay-off $\max(S_T - K, 0)$ and $\max(K - S_T, 0)$, respectively, at a given expiry day T . K is called strike price. In the traditional BS model it is assumed that the asset price process S_t follows a geometric Brownian motion with a constant diffusion coefficient σ . Under no arbitrage, the BS option pricing formula for calls is given by (Black and Scholes, 1973):

$$\begin{aligned} C_t^{BS}(S_t, K, \tau, r, \sigma) &= S_t \Phi(d_1) - K e^{-r\tau} \Phi(d_2), \\ d_1 &= \frac{\ln(S_t/K) + (r + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}, \\ d_2 &= d_1 - \sigma\sqrt{\tau}, \end{aligned} \tag{1}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of a standard normal random variable, r denotes the risk-free interest rate, S_t the price of underlying at time t , $\tau = T - t$ time to maturity and K the exercise price. Put prices are derived from the put-call parity

$$C_t - P_t = S_t - e^{-r\tau} K. \tag{2}$$

Empirically, the actual volatility σ of the underlying price process is the only parameter that cannot be observed directly. Hence one studies the volatility that is *implied* in option prices observed on markets, given the BS model were a true description of market conditions: implied volatility $\hat{\sigma}$ is defined as that σ which equates observed market prices \tilde{C}_t with the theoretical BS option price, i.e.

$$\hat{\sigma} : \quad C_t^{BS}(S_t, K, \tau, r, \hat{\sigma}) - \tilde{C}_t = 0. \tag{3}$$

BS formula is monotone in the volatility parameter σ , which ensures uniqueness of a solution $\hat{\sigma} > 0$. The purpose will be to estimate IVS given by $\hat{\sigma}_t : (K, \tau) \rightarrow \hat{\sigma}_t(K, \tau)$.

On the trading floors, only options near at-the-money, ATM, i.e. $K \approx S_t$ are traded at high liquidity. Since this region changes as the asset price fluctuates, it is convention to work on a moneyness metric, i.e. one scales strike prices by the current asset price (or future prices). Denote moneyness by $\kappa_t = K/S_t$. Then ATM is around $\kappa_t \approx 1$. A call option is called out-of-the-money, OTM, (in-of-the-money, ITM) if $\kappa > 1$ ($\kappa < 1$) with the reverse applying to puts.

Rewriting (1) in terms of moneyness yields

$$C_t^{BS}(S_t, K, \tau, r, \sigma) = S_t C(\kappa_t, \tau, r, \sigma)$$

where $C(\kappa_t, \tau, r, \sigma) = \Phi(d_1) - \frac{K}{S_t} e^{-r\tau} \Phi(d_2)$ and $d_1 = \frac{-\ln \kappa_t + (r + \frac{1}{2}\sigma^2)\tau}{\sigma\sqrt{\tau}}$, $d_2 = d_1 - \sigma\sqrt{\tau}$ (as before).

Our least squares kernel estimator is defined by

$$\hat{\sigma}(\kappa_t, \tau) = \arg \min_{\sigma} \sum_{i=1}^n \left\{ \tilde{C}_{t_i} - S_{t_i} C(\kappa_{t_i}, \tau_i, r_i, \sigma) \right\}^2 W(\kappa_{t_i}) K_1 \left(\frac{\kappa_t - \kappa_{t_i}}{h_{1,n}} \right) K_2 \left(\frac{\tau - \tau_i}{h_{2,n}} \right), \quad (4)$$

where $K_1(\cdot)$ and $K_2(\cdot)$ are kernel functions and $W(\cdot)$ is a weight function. $i = 1, \dots, n$ is a numbering of observed option prices.

We make the following assumptions:

(A1) $E\tilde{C}_t^4 < \infty$, $ES_t^4 < \infty$ and $EK^4 < \infty$, where E denotes the expectation operator.

(A2) $W(\cdot)$ is a uniformly continuous and bounded weight function.

(A3) $K_1(\cdot)$ and $K_2(\cdot)$ are bounded probability density kernel functions with bounded support.

(A4) Interest rate r is a fixed constant.

Assumption (A1), can be considered as a general characterization of the framework in which computing the implied volatility is ‘sensible’. Clearly, it holds in the BS model, but also

for the classes of diffusion and affine jump-diffusion models studied in finance, see e.g. Duffie et al. (2000). This is important, because practitioners like to use implied volatility in more general circumstances than the BS model provides. One may conjecture that $EK^4 < \infty$ is not needed because K is fixed in each option. However, when options are sampled randomly, it is useful to allow for this generalization. Note also that $E\tilde{C}_t^4 < \infty$ is actually an implication of $ES_t^4 < \infty$ and $EK^4 < \infty$, as long as we consider plain vanilla European options only. Since put options have a bounded payoff, there is no need for an additional assumption.

Assumption (A2) is very common, and some important weight function satisfy it. In Section (3) we will discuss possible choices of $W(\cdot)$. (A3) is a condition met by a lot of kernels used in nonparametric regression, such as the quartic or Epanechnikov kernel functions. (A4) is an assumption often used in the option pricing literature including the BS model. It is generally justified by the empirical observation that asset pricing variability largely outweighs changes in the interest rate. Nevertheless, the impact from changing interest rates can be substantial for options with a very long time to maturity. We also remark that including a fixed dividend yield into the model would be a straight forward extension of our estimator.

Given assumptions (A1) to (A4) we obtain consistency:

THEOREM 1. *Let $\sigma(\kappa_t, \tau)$ be the solution of $E[\{\tilde{C}_{t_1} - S_{t_1}C(\kappa_t, \tau, r_1, \sigma)\}W(\kappa_t)] = 0$. If conditions (A1), (A2), (A3) and (A4) are satisfied, we have*

$$\hat{\sigma}(\kappa_t, \tau) \xrightarrow{p} \sigma(\kappa_t, \tau)$$

as $nh_{1,n}h_{2,n} \rightarrow \infty$.

The proof is contained in the appendix. For our next result, we introduce the notations

$$\begin{aligned} A_i(\kappa_{t_i}, \tau_i, r_i, \sigma) &\stackrel{\text{def}}{=} \tilde{C}_{t_i} - S_{t_i}C(\kappa_{t_i}, \tau_i, r_i, \sigma) , \\ B(\kappa_t, \tau, r, \sigma) &\stackrel{\text{def}}{=} \frac{\partial C(\kappa_t, \tau, r, \sigma)}{\partial \sigma} = S_t^{-1} \frac{\partial C^{BS}(\cdot)}{\partial \sigma} = \sqrt{\tau} \phi(d_1) , \end{aligned} \quad (5)$$

$$D(\kappa_t, \tau, r, \sigma) \stackrel{\text{def}}{=} \frac{\partial^2 C(\kappa_t, \tau, r, \sigma)}{\partial^2 \sigma} = S_t^{-1} \frac{\partial^2 C^{BS}(\cdot)}{\partial \sigma^2} = \sqrt{\tau} \phi(d_1) d_1 d_2 \sigma^{-1} , \quad (6)$$

where $\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$. In financial language, B , the sensitivity of the option price with respect to implied volatility changes is called the option ‘vega’. It is an important quantity for portfolio managers wishing to hedge their volatility risk. The second derivative D is also

referred to as option ‘volga’. From the put-call parity in (2) it is seen that vega and volga are identical for calls and puts.

We establish

THEOREM 2. *Under conditions (A1), (A2), (A3), and (A4) if $E\{B^2(\kappa_t, \tau, r, \sigma)S_t^2W(\kappa_t)\} \neq E\{A(\kappa_t, \tau, r, \sigma)D(\kappa_t, \tau, r, \sigma)S_tW(\kappa_t)\}$, we have*

$$\sqrt{nh_{1n}h_{2n}}\{\hat{\sigma}(\kappa_t, \tau) - \sigma(\kappa_t, \tau)\} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \gamma^{-2}\nu^2),$$

where

$$\begin{aligned} \gamma^2 &\stackrel{\text{def}}{=} \left[E\{-B^2(\kappa_t, \tau, r, \sigma)S_t^2W(\kappa_t) + A(\kappa_t, \tau, r, \sigma)D(\kappa_t, \tau, r, \sigma)S_tW(\kappa_t)\} \right]^2 f(\kappa_t, \tau) \\ \nu^2 &\stackrel{\text{def}}{=} E\{A^2(\kappa_t, \tau, r, \sigma)B^2(\kappa_t, \tau, r, \sigma)S_t^2W^2(\kappa_t)\} \int K_1^2(u)K_2^2(v) dudv, \end{aligned}$$

and $f(\kappa_t, \tau)$ is the joint probability density function of κ_t and τ respectively.

The proof is again given in the appendix. Finally, the results carry over to put options:

COROLLAR 1. *By the put-call-parity and the bounded pay-off of put options, Theorem (1) and (2), hold also for put options, with A replaced correspondingly.*

The asymptotic distribution intricately depends on first and second order derivatives, and the particular weight function. Nevertheless an approximation is simple, since first and second order derivatives have the analytical expressions given in (5) and (6).

3 Applications

3.1 Weighting Functions, Kernels, and Minimization Scheme

In the vain of obtaining a good forecast of asset price variability, the early literature on implied volatility discusses weighting the observations intensively. Schmalensee and Trippi (1978) and Whaley (1982) argue in favor of unweighted averages, i.e. they use scalar estimate

$$\hat{\sigma}^* = \arg \min_{\sigma} \sum_{i=1}^n \{(\tilde{C}_i - C(\sigma))\}^2, \quad (7)$$

as a predictor of the future stock variability. Beckers (1981) minimizes

$$\hat{\sigma}^* = \arg \min_{\sigma} \sum_{i=1}^N w_i \{ \tilde{C}_i - C(\sigma) \}^2 / \sum_{i=1}^N w_i , \quad (8)$$

where $w_i \stackrel{\text{def}}{=} \partial C_i / \partial \sigma$, the option vega. Similarly, Latané and Rendelman (1976) use the squared vega as weights:

$$\hat{\sigma}^* = \sqrt{\sum_{i=1}^N w_i^2 \hat{\sigma}^2 / \sum_{i=1}^N w_i} , \quad (9)$$

Finally, Chiras and Manaster (1978) propose to employ the elasticity with respect to volatility:

$$\hat{\sigma}^* = \sum_{i=1}^N \eta_i \hat{\sigma} / \sum_{i=1}^N \eta_i , \quad (10)$$

where $\eta_i \stackrel{\text{def}}{=} \frac{\partial C_i}{\partial \sigma} \frac{\sigma}{C_i}$.

For calls and puts, vega is Gaussian shaped function in the underlying centered (roughly) ATM, Equation (5). Elasticity is a decreasing (increasing) function in the underlying for calls (puts). Common concern of the weighting procedures is to give low weight to ITM options, and highest weight to ATM or OTM options: ITM options are more expensive than ATM and OTM options because their intrinsic value, i.e. their payoff function evaluated at the current underlying prices, is already positive. Thus, they provide lower leverage for speculation, and produce higher costs in portfolio hedging. Due to their lower trading volume, they are suspected to sell at a liquidity premium which may ensue biased estimates of implied volatility. Consequently, some authors delete or downweigh ITM options, Aït-Sahalia and Lo (1998); Skiadopoulos et al. (1999).

Our estimator is general enough to allow for uniformly continuous and bounded weighting functions $W(\kappa)$ depending on moneyness. Technically, it is possible to use weights depending also on other variables including σ as done in (8) to (10). For several reasons, however, we like to refrain from using more involved weight functions: first, when ITM options are deleted or downweighted in the more recent literature, this choice is entirely determined by moneyness, not by vega. From this point of view, to have the weighting scheme depend on σ is rather implicit. Second, from a statistical point of view, weights depending on σ are likely to blow up the asymptotic variances in from of derivatives of W . This complicates estimation

and computation of confidence bands without adding to the problem of recovering a good estimate of the implied volatility surface. Finally, if one likes weights looking like the option vega or elasticity with respect to volatility, one may very easily construct weights $W(\kappa)$ that look very similar. For instance, an estimator in the type of Latané and Rendelman (1976) would put W shaped as a Gaussian density.

For the purpose of IVS estimation in our particular application, we like to give less weight to ITM options. This can be achieved by using as weighting functions

$$W(\kappa) = \frac{1}{\pi} \arctan\left\{\alpha(1 - \kappa)\right\} + 0.5 ,$$

for calls, and for puts

$$W(\kappa) = \frac{1}{\pi} \arctan\left\{\alpha(\kappa - 1)\right\} + 0.5 ,$$

where $\pi = 3.141\dots$ is the circle constant. α controls the speed, with which ITM options receive lower weight. ATM options are equally weighted. Outside $\kappa \approx 1$, only OTM options enter the minimization at significant weight. In our application we choose $\alpha = 9$. Other values are perfectly possible, and this choice is motivated to have a gentle transition between OTM call and OTM put options. The ultimate choice of α will depend on the specific application at hand.

The kernel functions employed throughout are quartic kernels, i.e.

$$K(u) = \frac{15}{16}(1 - u^2)^2 I(|u| \leq 1) ,$$

where $I(\mathcal{A})$ denotes the indicator function of the event \mathcal{A} . Other bounded kernels may be used, such as the Epanechnikov kernel. In practice, choice of the kernel functions has little impact on the estimates, Härdle (1990). Since the minimization is globally concave (compare proof of Theorem (1)), and well posed as long as h_1 and h_2 do not become unreasonably small, any minimization algorithm for globally concave objective functions can be employed. We use the Golden section search, Press et al. (1993), which is implemented in XploRe, www.i-xploRe.de. The tolerance, i.e. the fractional precision of the minimum, is maintained at 10^{-8} .

Observation date	Time to expiry (days)	min	max	mean	standard deviation	total number of observations	calls
Jan. 02, 2001	17	0.1711	0.3796	0.2450	0.0190	1219	561
	45	0.2112	0.2839	0.2425	0.0169	267	134
	73	0.1951	0.3190	0.2497	0.0199	391	209
	164	0.1777	0.3169	0.2528	0.0229	178	76
Feb. 02, 2001	14	0.1199	0.4615	0.1730	0.0211	1560	813
	42	0.1604	0.2858	0.1855	0.0188	715	329
	77	0.1628	0.2208	0.1910	0.0172	128	45
	133	0.1645	0.2457	0.1954	0.0221	119	63

Table 1: *Implied volatility data as obtained by inverting the BS formula separately for each observation in the sense of two-step estimators.*

3.2 Data Description and Estimations

The data used contains tick statistics on the DAX future and DAX index options and is provided by the German-Swiss Futures Exchange EUREX for January and February 2001. Both future and option data are contract based data, i.e. each single contract is registered together with its price, contract size, and time of settlement up to a hundredth second. Interest rate data in daily frequency, i.e. 1, 3, 6, 12 months EURIBOR rates, is gathered from *Thomson Financial Datastream*, and linearly interpolated to approximate the riskless interest rate for the option specific time to maturity.

For our application, we use data from January 02 and February 02, 2001. In a first step, we recover the DAX index values. To this end, we group to each option price observation C_t or P_t the future price F_t of the nearest available future contract, which was traded within a one minute interval around the observed option. The future price observation was taken from the most heavily traded future contract on the particular day, which is the March 2001 contract. The no-arbitrage price of the underlying index in a frictionless market without dividends is given by $S_t = F_t e^{-r_{T_F,t}(T_F-t)}$, where S_t and F_t denote the index and the future price respectively, T_F the future's maturity date, and $r_{T,t}$ the interest rate with maturity $T - t$.

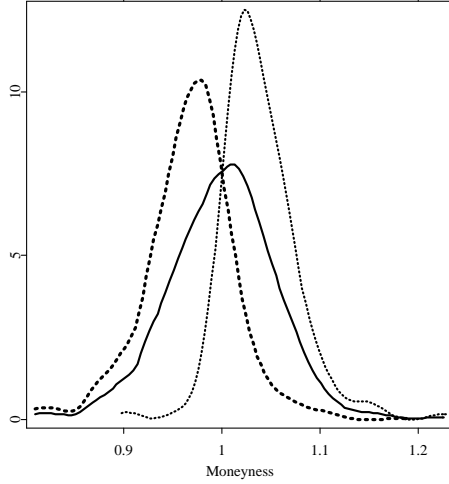


Figure 2: *Nonparametrically estimated densities of observed moneyness $\kappa_t = K/(F_t e^{-r\tau})$ for Jan. 02, 2001, and options with 17 days to expiry. Solid line for all observations, thickly dashed line for puts, and the more thinly dashed line for calls only. Quartic kernel used, bandwidth chosen according to Silverman’s rule of thumb. Density plots for other expiry dates including Feb. 02., 2001, available upon request.*

In the case of an capital weighted performance index as is the DAX index, Deutsche Börse (2002), dividends less corporate tax are reinvested into the index. Thus, dividends should not play any role for option valuation. In practice, however, one still observes – especially in late spring and early summer during the ‘dividend season’ of DAX index companies – deviances between put and call implied volatility, when the discounted futures are used as the underlying asset. For this case, a correction scheme is developed by Hafner and Wallmeier (2001), also described in Fengler et al. (2003b). In our case, this problem is mitigated by the fact that we use January and February data. Furthermore in using a weight function that downweights ITM options, which are most sensitive to this dividend wedge, this correction can also be achieved by the weighting function W itself.

In Table (1) we give an overview of the data employed. We prefer to present summary statistics in the form of implied volatility data obtained by inverting the BS formula separately for each observation rather than the option price data itself. The corresponding option data

can be seen in the top panel of Figure (3). Since the data is transaction data containing potential misprints, we also applied a filter in deleting all observations whose implied volatility is bigger than 0.7 and less than 0.1. For the distribution of the data across moneyness compare Figure (2) which presents density plots of moneyness for calls, puts, and all the observations observed on Jan.02 for 17 days to expiry. Put and call densities appear shifted which is due to the aforementioned liquidity argument of ATM and OTM options. For sake of space we do not present the very similar plots for the other expiry dates and Feb. 02, 2001. The settlement price of the March 2001 future was 6340 EUR at a volume of 30 088 contracts on Jan. 02, 2001, and 6669.5 EUR and 34 244 contracts on Feb. 02, 2001.

For our smile fits we pick the shortest time to expiry options of the Jan. 02 and Feb. 02, 2001, data. Plots are displayed in Figures 3 and 4. The top panel shows observed option prices given on the moneyness scale, while the lower panel demonstrates the estimate. In Figure 5, fits for the entire IVS are presented.

4 Conclusion

In this paper we present a new estimation approach to fit implied volatility smiles and surfaces. Usually one inverts first the Black-Scholes formula to back out implied volatility data and applies the fitting algorithm in a second step. This procedure may yield biased estimates when option price data is observed with measurement error only. Here, we provide a one-step estimator of the implied volatility surface. Our estimator is a least squares kernel estimator which directly smoothes implied volatility in the option price space. Asymptotic theory for the estimator is provided. We demonstrate the estimator using German DAX index option data to recover the smile and the implied volatility surface.

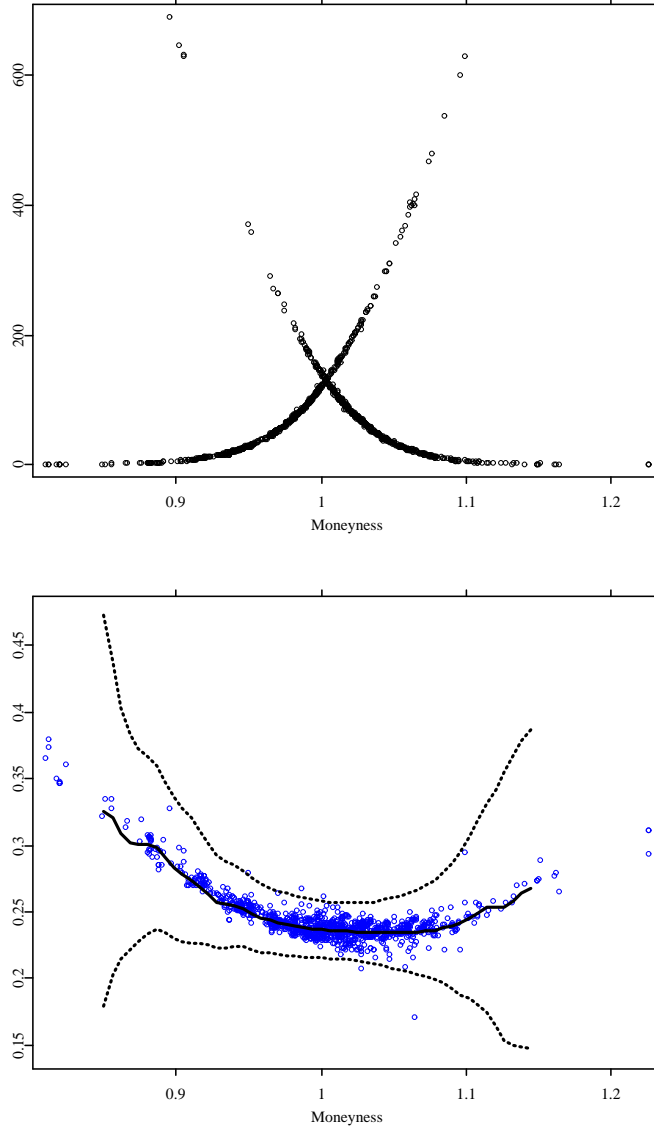


Figure 3: *Upper panel: Observed option price data on Jan 02, 2001. From lower left to upper right put prices, from upper left to lower right call prices. Lower panel: Least squares kernel smoothed implied volatility smile for 17 days to expiry on Jan 02, 2001. Bandwidth $h_1 = 0.025$, quartic Kernels employed. Minimization achieved by Golden section search. Dotted lines are the 95% confidence intervals for $\hat{\sigma}$. Single dots are implied volatility data obtained by inverting the BS formula separately for each observation in the sense of two-step estimators.*

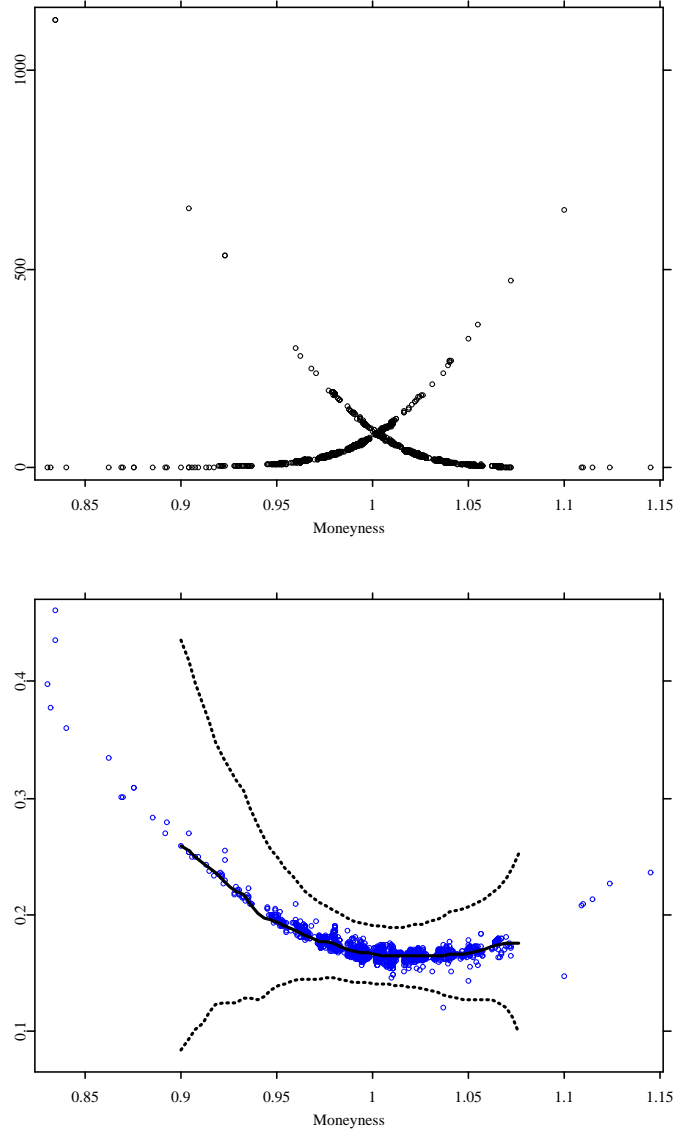
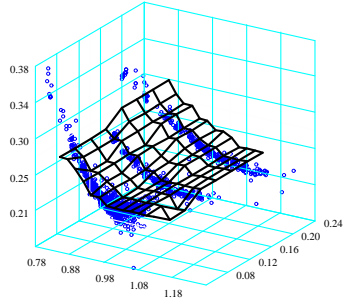


Figure 4: *Upper panel: Observed option price data on Feb. 02, 2001. From lower left to upper right put prices, from upper left to lower right call prices. Lower panel: Least squares kernel smoothed implied volatility smile for 14 days to expiry on Jan 02, 2001. Bandwidth $h_1 = 0.015$, quartic Kernels employed. Minimization achieved by Golden section search. Dotted lines are the 95% confidence intervals for $\hat{\sigma}$. Single dots are implied volatility data obtained by inverting the BS formula separately for each observation in the sense of two-step estimators.*

IVS on Jan. 02, 2001



IVS on Feb. 02, 2001

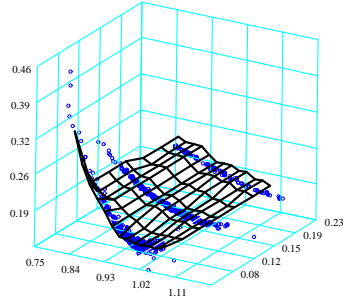


Figure 5: *Left panel: IVS fit for Jan. 02, 2001, with least squares kernel smoother. Bandwidths are $h_1 = 0.05$ in moneyness direction and $h_2 = 0.10$ in time to maturity direction. Right panel: IVS fit for Feb. 02, 2001, with least squares kernel smoother. Bandwidths are $h_1 = 0.03$ and $h_2 = 0.07$. In both panels, single dots denote implied volatility data obtained by inverting the BS formula separately for each observation in the sense of two-step estimators.*

References

- Aït-Sahalia, Y. and A. Lo**, “Nonparametric estimation of state-price densities implicit in financial asset prices,” *Journal of Finance*, 1998, *53*, 499–548.
- , **P. J. Bickel**, and **T. M. Stoker**, “Goodness-of-Fit Tests for Regression Using Kernel Methods,” *Journal of Econometrics*, 2001, *105*, 363–412.
- , **Y. Wang**, and **F. Yared**, “Do Options Markets Correctly Price the Probabilities of Movement of the Underlying Asset?,” *Journal of Econometrics*, 2001, *102*, 67–110.
- Ané, T. and H. Geman**, “Stochastic volatility and transaction time: an activity-based volatility estimator,” *Journal of Risk*, 1999, *2* (1), 57–69.
- Beckers, S.**, “Standard deviations implied in option prices as predictors of future stock price variability,” *Journal of Banking and Finance*, 1981, *5*, 363–382.
- Black, F. and M. Scholes**, “The pricing of options and corporate liabilities,” *Journal of Political Economy*, 1973, *81*, 637–654.
- Chiras, D. P. and S. Manaster**, “The information content of option prices and a test for market efficiency,” *Journal of Financial Economics*, 1978, *6*, 213–234.
- Cont, R. and J. da Fonseca**, “The Dynamics of Implied Volatility Surfaces,” *Quantitative Finance*, 2002, *2* (1), 45–602.
- Deutsche Börse**, *Leitfaden zu den Aktienindizes der Deutschen Börse*, 4.3 ed., Deutsche Börse AG, 2002.
- Duffie, D., J. Pan., and K. Singleton**, “Transform Analysis and Asset Pricing for Affine Jump-Diffusions,” *Econometrica*, 2000, *68*, 1343–1376.
- Dumas, B., J. Fleming, and R. E. Whaley**, “Implied Volatility Functions: Empirical Tests,” *Journal of Finance*, 1998, *LIII* (6), 2059–2106.
- Fengler, M. R., W. Härdle, and C. Villa**, “The dynamics of implied volatilities: A common principle components analysis,” Discussion Paper 2001-38, SFB 373, Humboldt-University, Berlin 2003.

- , – , and **E. Mammen**, “Implied Volatility String Dynamics,” Sfb 373 Discussion Paper, Humboldt-Universität zu Berlin 2003.
- , – , and **P. Schmidt**, “Common Factors Governing VDAX Movements and the Maximum Loss,” *Journal of Financial Markets and Portfolio Management*, 2002, 16 (1), 16–29.
- Hafner, R. and M. Wallmeier**, “The Dynamics of DAX Implied Volatilities,” *International Quarterly Journal of Finance*, 2001, 1 (1), 1–27.
- Härdle, W.**, *Applied Nonparametric Regression*, Cambridge, UK: Cambridge University Press, 1990. www.i-xplore.de.
- Harvey, C. R. and R. E. Whaley**, “S&P 100 Index Option Volatility,” *Journal of Finance*, 1991, 46 (4), 1151–1561.
- Hentschel, L.**, “Errors in implied volatility estimation,” *Journal of Financial and Quantitative Analysis*, 2002. Forthcoming.
- Latané, H. A. and J. Rendelman**, “Standard deviations of stock price ratios implied in option prices,” *Journal of Finance*, 1976, 31, 369–381.
- Manaster, S. and G. Koehler**, “The calculation of implied variances from the Black-and-Scholes model: A note,” *Journal of Finance*, 1982, 37, 227–230.
- Press, W., B. Flannery, S. Teukolsky, and W. Vetterling**, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, 1993.
- Roll, R.**, “A simple implicit measure of the effective bid-ask spread,” *Journal of Finance*, 1984, 39, 1127–1139.
- Rookley, C.**, “Fully Exploiting the Information Content of Intra-Day Option Quotes: Applications in Option Pricing and Risk Management,” Technical Report, Department of Finance, University of Arizona 1997.
- Rosenberg, J.**, “Implied volatility functions: A reprise,” *Journal of Derivatives*, 2000, 7, 51–64.
- Schmalensee, R. and R. R. Trippi**, “Common Stock Volatility Expectations Implied by Option Premia,” *Journal of Finance*, 1978, 33, 129–147.

Shimko, D., “Bounds on probability,” *Risk*, 1993, *6*, 33–37.

Skiadopoulos, G., S. Hodges, and L. Clewlow, “The Dynamics of the S&P 500 Implied Volatility Surface,” *Review of Derivatives Research*, 1999, *3*, 263–282.

Tompkins, R., “Implied volatility surfaces: Uncovering regularities for options on financial futures,” Working Paper, Vienna University of Technology 1999.

Whaley, R., “Valuation of american call options on dividend-paying stocks: empirical tests,” *Journal of Financial Economics*, 1982, *10*, 29–58.

A Proofs

PROOF OF THEOREM (1):

Let

$$\widehat{L}_n(\sigma) \stackrel{\text{def}}{=} \frac{1}{nh_{1,n}h_{2,n}} \sum_{i=1}^n \{\tilde{C}_{t_i} - S_{t_i}C(\kappa_{t_i}, \tau_i, r_i, \sigma)\}^2 W(\kappa_{t_i}) K_1\left(\frac{\kappa_t - \kappa_{t_i}}{h_{1,n}}\right) K_2\left(\frac{\tau - \tau_i}{h_{2,n}}\right).$$

As a first step, let us prove

$$\widehat{L}_n(\sigma) \xrightarrow{p} L(\sigma) \stackrel{\text{def}}{=} E\left[\{\tilde{C}_t - S_t C(\kappa_t, \tau, r; \sigma)\}^2 W(\kappa_t)\right]. \quad (11)$$

It is observed that

$$\begin{aligned} \widehat{L}_n(\sigma) &= \frac{1}{nh_{1,n}h_{2,n}} \sum_{i=1}^n \left\{ \{\tilde{C}_{t_i} - S_{t_i}C(\kappa_{t_i}, \tau_i, r_i; \sigma)\}^2 W(\kappa_{t_i}) K_1\left(\frac{\kappa_t - \kappa_{t_i}}{h_{1,n}}\right) K_2\left(\frac{\tau - \tau_i}{h_{2,n}}\right) \right. \\ &\quad \left. - E\left[\{\tilde{C}_{t_i} - S_{t_i}C(\kappa_{t_i}, \tau_i, r_i; \sigma)\}^2 W(\kappa_{t_i}) K_1\left(\frac{\kappa_t - \kappa_{t_i}}{h_{1,n}}\right) K_2\left(\frac{\tau - \tau_i}{h_{2,n}}\right)\right] \right\} \\ &\quad + \frac{1}{h_{1,n}h_{2,n}} E\left[\{\tilde{C}_{t_1} - S_{t_1}C(\kappa_{t_1}, \tau_1, r_1; \sigma)\}^2 W(\kappa_{t_1}) K_1\left(\frac{\kappa_t - \kappa_{t_1}}{h_{1,n}}\right) K_2\left(\frac{\tau - \tau_1}{h_{2,n}}\right)\right] \\ &\stackrel{\text{def}}{=} \alpha_n + \beta_n \end{aligned} \quad (12)$$

Standard arguments can be used to prove

$$E\alpha_n^2 = \mathcal{O}((nh_{1,n}h_{2,n})^{-1}) \quad (13)$$

by conditions (A1) and (A2).

By Taylor's expansion, we have

$$\begin{aligned} \beta_n &= \frac{1}{h_{n,1}h_{n,2}} E \int \{\tilde{C}_{t_1} - S_{t_1}C(x, y, r_1; \sigma)\}^2 W(x) K_1\left(\frac{\kappa_t - x}{h_{1,n}}\right) K_2\left(\frac{\tau - y}{h_{2,n}}\right) dx dy \\ &= E \int \{\tilde{C}_t - S_t C(\kappa_t - h_n u, \tau - h_n v, r; \sigma)\}^2 W(\kappa_t - h_n u) K_1(u) K_2(v) du dv \\ &\xrightarrow{p} L(\sigma). \end{aligned} \quad (14)$$

(13) and (14) together prove (11).

In a second step, we have, recalling the definition of $\sigma(\kappa_t, \tau)$

$$\begin{aligned} \frac{\partial L(\sigma)}{\partial \sigma} \Big|_{\sigma=\sigma(\kappa_t, \tau)} &= -2E\tilde{C}_t S_t W(\kappa_t) \frac{\partial}{\partial \sigma} C(\kappa_t, \tau, r, \sigma) \Big|_{\sigma=\sigma(\kappa_t, \tau)} \\ &\quad + 2ES_t^2 C(\kappa_t, \tau, r, \sigma(\kappa_t, \tau)) W(\kappa_t) \frac{\partial}{\partial \sigma} C(\kappa_t, \tau, r, \sigma) \Big|_{\sigma=\sigma(\kappa_t, \tau)} = 0 \end{aligned}$$

and

$$\begin{aligned} &\frac{\partial^2 L(\sigma)}{\partial \sigma^2} \Big|_{\sigma=\sigma(\kappa_t, \tau)} \\ &= -2E\tilde{C}_t S_t W(\kappa_t) \frac{\partial^2}{\partial \sigma^2} C(\kappa_t, \tau, r, \sigma) \Big|_{\sigma=\sigma(\kappa_t, \tau)} \\ &\quad + 2ES_t^2 W(\kappa_t) \left(\frac{\partial}{\partial \sigma} C(\kappa_t, \tau, r, \sigma) \Big|_{\sigma=\sigma(\kappa_t, \tau)} \right)^2 \\ &\quad + 2ES_t^2 W(\kappa_t) C(\kappa_t, \tau, r, \sigma(\kappa_t, \tau)) \frac{\partial^2}{\partial \sigma^2} C(\kappa_t, \tau, r, \sigma) \Big|_{\sigma=\sigma(\kappa_t, \tau)} \\ &= 2ES_t^2 W(\kappa_t) \left(\frac{\partial}{\partial \sigma} C(\kappa_t, \tau, r, \sigma) \Big|_{\sigma=\sigma(\kappa_t, \tau)} \right)^2 . \end{aligned}$$

This together with (11) proves that $\hat{L}_n(\sigma)$ converges in probability to a concave function with a unique minimum at $\sigma = \sigma(\kappa_t, \tau)$. Thus $\hat{\sigma}_n(\kappa_t, \tau) \xrightarrow{p} \sigma(\kappa_t, \tau)$ is proved.

PROOF OF THEOREM (2):

Recalling the definition of $\hat{\sigma}(\kappa_t, \tau)$, it follows that $\hat{\sigma}(\kappa_t, \tau)$ is the solution of the following equation

$$\begin{aligned} U_n(\sigma) &\stackrel{\text{def}}{=} (nh_{1n}h_{2n})^{-1} \sum_{i=1}^n S_{t_i} A_i(\kappa_{t_i}, \tau_i, r_i, \sigma) B_i(\kappa_{t_i}, \tau_i, r_i, \sigma) W(\kappa_{t_i}) K_1 \left(\frac{\kappa_t - \kappa_{t_i}}{h_{1,n}} \right) K_2 \left(\frac{\tau - \tau_i}{h_{2,n}} \right) \\ &= 0 . \end{aligned}$$

By Taylor's expansion, we get

$$0 = U_n(\hat{\sigma}(\kappa_t, \tau)) = U_n(\sigma(\kappa_t, \tau)) + U'_n(\sigma^*)(\hat{\sigma}_t(\kappa_t, \tau) - \sigma(\kappa_t, \tau)), \quad (15)$$

where σ^* lies between σ and $\hat{\sigma}$ and $U'_n(\sigma^*) \stackrel{\text{def}}{=} \frac{\partial}{\partial \sigma} U_n(\sigma) \Big|_{\sigma=\sigma^*}$. By (15), we have

$$\hat{\sigma}(\kappa_t, \tau) - \sigma(\kappa_t, \tau) = -[U'_n(\sigma^*)]^{-1} U_n(\sigma) . \quad (16)$$

By some algebra, we obtain

$$\begin{aligned}
U'_n(\sigma) &= \frac{1}{nh_{1,n}h_{2,n}} \sum_{i=1}^n \left\{ \left[\left(\frac{\partial}{\partial \sigma} A_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \right) B_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \right. \right. \\
&\quad \left. \left. + A_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \left(\frac{\partial}{\partial \sigma} B_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \right) \right] S_{t_i} W(\kappa_{t_i}) \right. \\
&\quad \times K_1 \left(\frac{\kappa_t - \kappa_{t_i}}{h_n} \right) K_2 \left(\frac{\tau - \tau_i}{h_{2n}} \right) \\
&\quad \left. - E \left[\left(\frac{\partial}{\partial \sigma} A_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \right) B_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \right. \right. \\
&\quad \left. \left. + A_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \frac{\partial}{\partial \sigma} B_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \right] S_{t_i} W(\kappa_{t_i}) \right. \\
&\quad \left. \times K_1 \left(\frac{\kappa_t - \kappa_{t_i}}{h_{1,n}} \right) K_2 \left(\frac{\tau - \tau_i}{h_{2,n}} \right) \right\} \\
&\quad + \frac{1}{nh_{1n}h_{2n}} \sum_{i=1}^n E \left\{ \left[\left(\frac{\partial}{\partial \sigma} A_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \right) B_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \right. \right. \\
&\quad \left. \left. + A_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \frac{\partial}{\partial \sigma} B_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \right] S_{t_i} W(\kappa_{t_i}) \right. \\
&\quad \left. \times K_1 \left(\frac{\kappa_t - \kappa_{t_i}}{h_n} \right) K_2 \left(\frac{\tau - \tau_i}{h_{2n}} \right) \right\} \\
&\stackrel{\text{def}}{=} \Delta_{n1} + \Delta_{n2}
\end{aligned} \tag{17}$$

Inspect first Δ_{n1} in Equation (18): by some algebra, we get

$$\begin{aligned}
&E \Delta_{n1}^2 \\
&\leq \frac{1}{n^2 h_{1,n}^2 h_{2,n}^2} \sum_{i=1}^n E \left\{ \left[\left(\frac{\partial}{\partial \sigma} A_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \right) B_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \right. \right. \\
&\quad \left. \left. + A_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \frac{\partial}{\partial \sigma} B_i(\kappa_{t_i}, \tau_i, r_i, \sigma) \right] S_{t_i} W(\kappa_{t_i}) \right. \\
&\quad \left. \times K_1 \left(\frac{\kappa_t - \kappa_{t_i}}{h_n} \right) K_2 \left(\frac{\tau - \tau_i}{h_{2n}} \right) \right\}^2 \\
&= \frac{f^2(\kappa_t, \tau) \int K_1^2(u) du \int K_2^2(v) dv}{nh_{1,n}h_{2,n}} E \left\{ \left[\left(\frac{\partial}{\partial \sigma} A_1(\kappa_t, \tau, r_1, \sigma) B_1(\kappa_t, \tau, r_1, \sigma) \right. \right. \right. \\
&\quad \left. \left. \left. + A_1(\kappa_t, \tau, r_1, \sigma) \frac{\partial}{\partial \sigma} B_1(\kappa_t, \tau, r_1, \sigma) \right)^2 \right] S_t^2 W(\kappa_t) \right\} + o \left(\frac{1}{nh_{1,n}h_{2,n}} \right) \longrightarrow 0. \tag{19}
\end{aligned}$$

as $nh_{1,n}h_{2,n} \rightarrow \infty$. $f(\kappa_t, \tau)$ denotes the joint probability density functions of κ_t and τ .

To consider Δ_{n2} in Equation (18), denote $D(\kappa_t, \tau, r, \sigma) = \frac{\partial}{\partial \sigma} B(\kappa_t, \tau, r, \sigma)$, for simplicity. Note that $\frac{\partial}{\partial \sigma} A(\kappa_t, \tau, r, \sigma) = -S_t B(\kappa_t, \tau, r, \sigma)$. Thus, we have

$$\begin{aligned}
\Delta_{n2} &= \frac{1}{h_{1n}h_{2n}} E \left[\int \left(-B^2(x, y, r, \sigma) + A(x, y, r, \sigma) D(x, y, r_1, \sigma) \right) S_t W(\kappa_t) \right. \\
&\quad \times K_1 \left(\frac{\kappa_t - x}{h_{1,n}} \right) K_2 \left(\frac{\kappa_t - x}{h_{2,n}} \right) f(x, y) dx dy \Big] \\
&= E \int [-B^2(\kappa_t - h_{1,n}u, \tau - h_{2,n}v, r_1, \sigma) S_t \\
&\quad + A(\kappa_t - h_{1,n}u, \tau - h_{2,n}v, r_1, \sigma) D(\kappa_t - h_{1,n}u, \tau - h_{2,n}v, r_1, \sigma)] S_t W(\kappa_t) \\
&\quad \times f_1(\kappa_t - h_{1,n}u, \tau - h_{2,n}v) K_1(u) K_2(v) du dv \\
&\longrightarrow \left\{ -E[B^2(\kappa_t, \tau, r, \sigma) S_t^2 W(\kappa_t)] + E[A(\kappa, \tau, r_1, \sigma) D(\kappa, \tau, r_1, \sigma) S_t W(\kappa_t)] \right\} f(\kappa_t, \tau)
\end{aligned} \tag{20}$$

(18), (19), (20) and the fact $U'_n(\sigma^*) - U'_n(\sigma) \rightarrow 0$ together prove

$$U'_n(\sigma^*) \xrightarrow{p} \left\{ E[-B^2(\kappa_t, \tau, r, \sigma) S_t^2 W(\kappa_t)] + E[A(\kappa, \tau, r, \sigma) D(\kappa, \tau, r, \sigma) S_t W(\kappa_t)] \right\} f(\kappa_t, \tau). \tag{21}$$

Now, let

$$u_{ni} \stackrel{\text{def}}{=} \frac{1}{h_{1,n}h_{2,n}} A(\kappa_{t_i}, \tau_i, r_i, \sigma) B(\kappa_{t_i}, \tau_i, r_i, \sigma) S_{t_i} W(\kappa_{t_i}) K_1 \left(\frac{\kappa - \kappa_{t_i}}{h_{1,n}} \right) K_2 \left(\frac{\tau - \tau_i}{h_{2,n}} \right).$$

For some $\delta > 0$, we have

$$\begin{aligned}
E|u_{ni}|^{2+\delta} &= \frac{1}{h_{1,n}^{2+\delta} h_{2,n}^{2+\delta}} E A^{2+\delta}(\kappa_{t_i}, \tau_i, r_i, \sigma) B^{2+\delta}(\kappa_{t_i}, \tau_i, r_i, \sigma) S_{t_i}^{2+\delta} W^{2+\delta}(\kappa_{t_i}) \\
&\quad \times K_1^{2+\delta} \left(\frac{\kappa_t - \kappa_{t_i}}{h_{1,n}} \right) K_2^{2+\delta} \left(\frac{\tau - \tau_i}{h_{2,n}} \right) \\
&= \frac{1}{h_{1,n}^{1+\delta} h_{2,n}^{1+\delta}} E \int A^{2+\delta}(\kappa_t - h_n u, \tau - h_n v, r, \sigma) B^{2+\delta}(\kappa_t - h_n u, \tau - h_n v, r, \sigma) \\
&\quad \times S_t^{2+\delta} W^{2+\delta}(\kappa_t - h_n u) K_1^{2+\delta}(u) K_2^{2+\delta}(v) du dv \\
&= \frac{f(\kappa_t, \tau) \int K_1^{2+\delta}(u) du \int K_2^{2+\delta}(v) dv}{h_{1,n}^{1+\delta} h_{2,n}^{2+\delta}} E [A^{2+\delta}(\kappa_t, \tau, \sigma) B^{2+\delta}(\kappa_t, \tau, \sigma) S_t^{2+\delta} W^{2+\delta}(\kappa_t)] \\
&\quad + o \left(\frac{1}{h_{1,n}^{1+\delta} h_{2,n}^{1+\delta}} \right).
\end{aligned} \tag{22}$$

Similarly, we get

$$\begin{aligned}
Eu_{ni}^2 &= \frac{f(\kappa_t, \tau) \int K_1^2(u) du \int K_2^2(v) dv}{h_{1,n} h_{2,n}} E\{A^2(\kappa_t, \tau, r, \sigma) B^2(\kappa, \tau, r, \sigma) S_t^2 W^2(\kappa_t)\} \\
&\quad + \mathcal{O}\left(\frac{1}{h_{1,n} h_{2,n}}\right).
\end{aligned} \tag{23}$$

(22) and (23) together prove

$$\frac{\sum_{i=1}^n E|u_{ni}|^{2+\delta}}{(\sum_{i=1}^n E|u_{ni}|^2)^{\frac{2+\delta}{2}}} = \mathcal{O}((nh_{1,n}h_{2,n})^{-\frac{\delta}{2}}) = \mathcal{O}(1)$$

as $nh_{1,n}h_{2,n} \rightarrow 0$.

Applying the Liapounov central limit theorem, we get

$$\sqrt{nh_{1,n}h_{2,n}} U_n(\sigma) \xrightarrow{\mathcal{L}} \mathcal{N}(0, f(\kappa_t, \tau) \nu^2), \tag{24}$$

where ν^2 is defined in Theorem (2).

By (21) and (24), Theorem (2) is proved.